Kramer, E. M., Dorit, R. L., and Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development: Duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* **149,** 765–783.

Ruiz-Medrano, R., Xoconostle-Cazares, B., and Lucsa, W. (1999). Phloem long-distance transport of CmNACP mRNA: Implications for supracellular regulation in plants. *Development* **126,** 4405–4419.

Sambrook, J., and Russel, D. W. (2001a). ''Molecular Cloning,'' Vol. 2. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Sambrook, J., and Russel, D. W. (2001b). ''Molecular Cloning,'' Vol. 3. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Suzuki, Y., Hibino, T., Kawazu, T., Wada, T., Kihara, T., and Koyama, H. (2003). Extraction of total RNA from leaves of Eucalyptus and other woody and herbaceous plants using sodium isoascorbate. *Biotechniques* **34,** 988–993.

Wang, S. X., Hunter, W., and Plant, A. (2000). Isolation and purification of functional total RNA form woody branches and needles of Sitka and white spruce. *Biotechniques* **28,** 292–296.

Weigel, D., and Glazebrook, J. (2002). ''Arabidopsis: A Laboratory Manual.'' Cold Spring Harbor Press, Cold Spring Harbor, NY.

Zachgo, S., Perbal, M.-C., Saedler, H., and Schwarz-Sommer, Z. (2000). *In situ* analysis of RNA and protein expression in whole mounts facilitates detection of floral gene expression dynamics. *Plant J.* **23,** 697–702.

[33] Developing Antibodies to Synthetic Peptides Based on Comparative DNA Sequencing of Multigene Families

By ROGER H. SAWYER, TRAVIS C. GLENN, JEFFREY O. FRENCH, and LOREN W. KNAPP

Abstract

Using antisera to analyze the expression of specific gene products is a common procedure. However, in multigene families, such as the β-keratins of the avian integument where strong homology exists among the scale (ScβK), claw (ClβK), feather (FβK), and feather-like (FlβK) subfamilies, determining the cellular and tissue expression patterns of the subfamilies is difficult because polyclonal antisera produced from any one protein recognize all family members. Traditionally, researchers produced and screened multiple monoclonal antisera produced from the proteins of interest until an antiserum with sufficient specificity could be obtained. Unfortunately, this approach requires a lot of effort, and once obtained, such antisera may have limited applications. Here, we present procedures by which comparative DNA sequences of members from the β-keratin multigene family were

0076-6879/05 $35.00

translated and aligned to identify amino acid domains that were conserved within the FβK subfamily, but which were divergent from the other subfamilies. A synthetic 23-mer peptide with the conserved amino acid sequence was generated and used to produce a polyclonal antiserum that recognizes only the FβK subfamily of proteins. Western blot analysis and confocal microscopy with this antiserum are now providing valuable new insights concerning the developmental and evolutionary relationships between the scale, claw, and feather proteins found in birds. This represents a powerful new approach combining techniques from molecular evolution and developmental biology to study the expression and evolution of specific members of multigene families.

Introduction

Functional genomics encompasses many aspects of biology ranging from highly experimental systems in which variation in transcription and translation are monitored closely *in vitro*, to overarching relationships in regulation of multigene families across taxa of organisms with shared evolutionary histories (http://www.nsf.gov/pubs/2001/bio012/). Here, we present techniques that have allowed us to address the role of a multigene family in the evolution of a phenotypic innovation, namely the avian feather (Sawyer *et al*., 2003a). Our model examines the role of the multigene family of epidermal β-keratins whose expression is restricted to the skin and its appendages in reptiles and birds (Alibardi and Sawyer, 2002; Sawyer *et al*., 2000, 2003b). The expression of this family of genes is intimately associated with the morphogenesis of integumentary appendages (Haake *et al*., 1984; Sawyer and Knapp, 2003), so these unique and highly insoluble protein products (Shames *et al*., 1988) make up complex structural elements, such as the barbs, barbules, and hooklets of feathers (Haake *et al*., 1984). Understanding the evolutionary origin of feathers depends on understanding the evolution of development, including an understanding of the evolution of the developmental expression of the members of the β-keratin gene family.

The overall approach that we describe is generic, uses common techniques, and can be applied to study many multigene families. Many of the specific methods that we have used draw on techniques from diverse fields of biology, and thus, the salient points of all the methods chosen may not be readily apparent to many readers. In this chapter, we describe the rationale for each method and the important goals that must be achieved and illustrate the method with specific techniques we have used to study β-keratins. By describing methods in this way, we hope that the overall approach will be accessible to a larger pool of developmental and evolutionary biologists.

Background on Example Used

Four subfamilies of β-keratins are known for the chicken: scale β-keratins (ScβK), claw β-keratins (ClβK), feather β-keratins (FβK), and feather-like β-keratins (FlβK) (Presland *et al*., 1989a,b; Shames *et al*., 1988). Because of their poor solubility, tendency to aggregate in gels, and complex expression patterns in embryonic epidermal appendages, generation of non–cross-reacting antisera has not been possible (Shames *et al*., 1988). To date, polyclonal antisera generated against the ScβK isolated by one- and two-dimensional polyacrylamide gel electrophoresis have cross-reacted with the FβK on Western blots and with feather tissues (Carver and Sawyer, 1987; Haake *et al*., 1984; O'Guin *et al*., 1982; Sawyer *et al*., 2000; Shames *et al*., 1988). These antisera, such as ''anti-β_1'' (Sawyer *et al*., 2000; Shames *et al*., 1988) also react with the β-keratins found in the epidermal appendage of reptiles (Alibardi and Sawyer, 2002).

Here, we describe a comparative genomic approach, in which we have taken advantage of DNA and protein sequence data in GenBank to align the four subfamilies of the β-keratin family and design polymerase chain reaction (PCR) primers specific for the FβK subfamily. By aligning the inferred amino acid sequences of the FβK from several orders of birds, it is possible to identify highly conserved peptide sequences within the FβK subfamily, which are divergent from the other subfamilies. A synthetic peptide with the identified sequence can then be used to generate specific antisera. Such antisera do not cross-react with the other subfamilies of the β-keratin multigene family, thus allowing specific localization of the FβK subfamilies of proteins in other skin appendages of birds and possibly reptiles (Sawyer *et al*., 2003a,b). Antisera generated in such a manner may also be used to analyze the expression of the FβKs during the development of various skin appendages using both Western blots of extracted proteins and indirect immunofluorescence of tissues (Sawyer *et al*., 2003a,b).

Methods (Fig. 1)

Initial Characterization of the Multigene Family

Information on the expression of multigene families during embryogenesis has generally come from comparative analysis of proteins extracted from stage-specific tissues (Kemp and Rogers, 1972; Molloy *et al*., 1982; O'Guin and Sawyer, 1982). Direct amino acid sequencing of such proteins provides preliminary sequence information (Inglis *et al*., 1987; Walker and Bridgen, 1976). The generation of cDNA libraries from messenger RNA (mRNA)

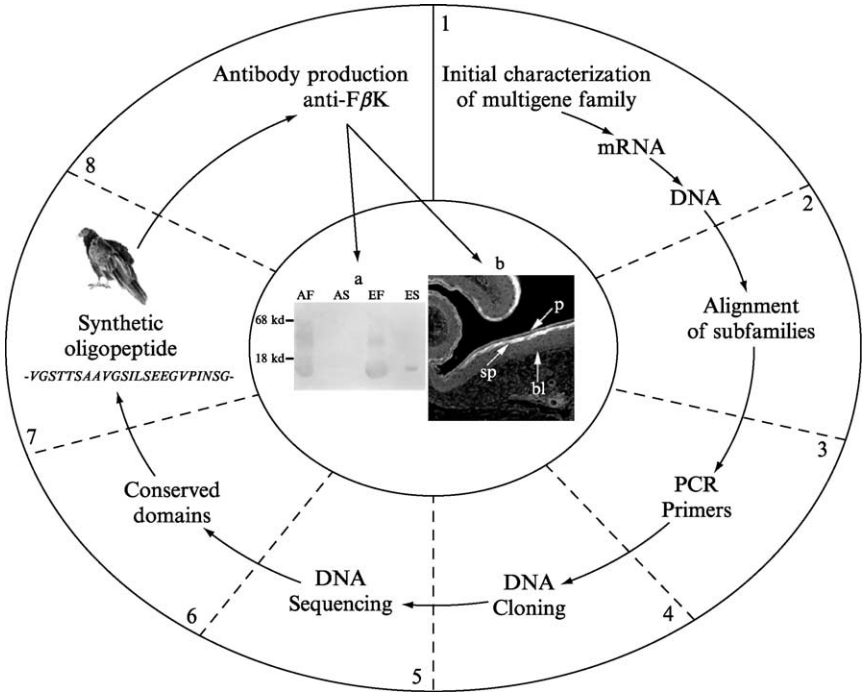


FIG. 1. A schematic representation of the procedures used to obtain antisera against specific conserved peptides, identified through genomic analysis. Each number around the circle, 1–8, represents a step in the procedure and corresponds to the eight headings in this manuscript. The letters *a* and *b* in Fig. 1 show a Western blot and a confocal image, respectively. These figures are described in detail in Figs. 3 and 4, respectively.

populations provides the initial sequence data (Presland *et al*., 1989a,b; Wilton *et al*., 1985), and analysis of genomic clones then provides comparative information on gene sequence and organization (Presland *et al*., 1989a,b; Rogers *et al*., 1998). Hopefully these data provide enough sequence information so that conserved regions can be identified and PCR primers can be designed.

Example, β-*Keratins (Fig. 1)*

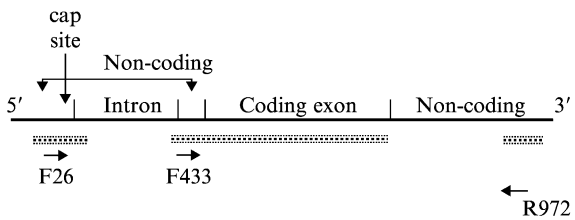
Initially, the *β*-keratins of scales and feathers were analyzed by gel electrophoresis (Kemp and Rogers, 1972). Using direct sequencing of polypeptides, the amino acid sequences of a few *β*-keratins were published (Inglis *et al*., 1987; Walker and Bridgen, 1976), yet it was not until cDNAs were generated from mRNA populations that DNA sequences for the F*β*K

and ScβK were first described (Presland *et al*., 1989a,b; Wilton *et al*., 1985). Subsequently, using feather and scale cDNAs as probes, genomic libraries provided information on the genomic organization of the subfamilies of β-keratins (Molloy *et al*., 1982; Presland *et al*., 1989a,b; Rogers *et al*., 1998; Whitbread *et al*., 1991; Wilton *et al*., 1985). The overall structure of the FβKs is presented in Fig. 2.

If a sufficient amount of information exists (i.e., can be found in GenBank), then steps 2–5 may be omitted. Here, we assume that a small, but limited, amount of initial information is available. If that is true, then it will be often desirable to increase the amount of comparative information available to identify conserved domains before attempting to design the oligopeptides (step 7). By using well-developed and widely available tools from molecular genetics and molecular evolution (steps 2–5), researchers can increase the amount of comparative information available relatively quickly and easily. In most situations, researchers will want to skip ahead to step 6, determine how much additional information is desired, and then come back to steps 2–5.

Aligning DNA Sequences

After obtaining DNA sequence information for some members of the multigene family, the next step is to align the sequences. A variety of



Primer	Sequence (5′→3′)
F26	CGC CCT CAT CCA CKT CTC TT
F433	CCA TGT CCT GCT ATG ATC TGT G
R972	CTC AAC TTG CTT CAG GAT YAA

FIG. 2. Feather β-keratin gene structure (summarizing Presland *et al*. [1989a]) and the locations of the primers used in polymerase chain raction (PCR) and sequencing. F26 and R972 were used in all successful PCRs. F433 and R972 were used in sequencing of the focused dataset (after Sawyer *et al*. [2003]).

software tools and protocols are available to aid in this process. One of the best and mostly widely used alignment tools is Clustal (Thompson *et al*., 1994, 1997). Stand alone versions of ClustalX are freely available (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/), although some researchers may benefit from step-by-step guides (e.g., Hall, 2001; http://www.biozentrum.unibas.ch/~biophit/clustal/ClustalX_help.html). The Clustal algorithm is also incorporated into a variety of commercial (GCG, DS Gene, MacVector, etc.) and free (http://workbench.sdsc.edu/) software programs and web sites. Researchers with little experience in this area should keep in mind that software packages can be very sensitive to input parameters and variation in sequence length, especially inconsistencies in the homology of beginning and ending locations. Because the alignment produced in this situation is used simply to identify conserved regions of DNA sequence to be used in the subsequent steps of this protocol, many researchers may find manual alignment the quickest way to achieve this goal.

Example, the β*-Keratin Multigene Family (Fig. 1).* It is desirable to identify conserved sequences that will allow amplification of the entire coding region, especially when the coding region is small, as is the case for FβKs. Presland *et al*. (1989a) include alignments of the β-keratin data available at the time of publication. Initially, we supplemented the published alignments using GCG (Accelrys, Inc., San Diego, CA) to align all sequences of β-keratin available. Subsequently, we have used the ClustalW algorithm (Thompson *et al*., 1994), implemented in Sequence Navigator 1.0.1 (Applied Biosystems). From these alignments, it was clear that a region including the cap site, 5′ to the coding region, and a 3′ untranslated region contained few substitutions among the sequences available.

Designing and Using PCR Primers

After identifying conserved regions of DNA sequence, primers for the polymerase chain reaction (PCR) need to be designed. Again, various software tools and protocols (Sambrook and Russell, 2001) are available to aid in this process. One commercial software package (Oligo 6 [Molecular Biology Insights, Cascade, CO]) includes an option to identify primer-binding regions from alignments of multiple sequences. Most researchers, however, will find it easier and more economical to simply import the consensus sequence or a representative sequence into the chosen primer design software. Primers can then be designed from conserved regions of the DNA sequence. If variable sites occur in potential primer-binding regions, then mixed bases can be integrated into the primers (Palumbi, 1996). Use of mixed bases should be minimized, especially on the 3′ ends of primers. Many software packages do not allow mixed bases to be analyzed

automatically, but they will allow primers to be manually edited. By manually editing potential primers, the effects of different specific bases at variable sites can be investigated.

Example, Feather β*-keratin Subfamily (Fig. 1).* Using aligned feather *β*-keratin DNA sequences from chicken (Presland *et al*., 1989a,b), primers for PCR were designed in two conserved noncoding regions, which span the entire coding region (Fig. 2). These primers were tested by using them in a PCR with chicken genomic DNA as the template. Primers F26 and R972 successfully amplify about 900–1000 bp fragments of DNA from a variety of birds (French, 2001).

Cloning Members of the Multigene Family

Because multiple members of the gene family are amplified during PCR, it is not possible to determine the sequences directly from the PCR products. First, the PCR products must be separated so that amplicons from individual members can be sequenced separately. The most straightforward way to accomplish this is to clone the PCR products. During the process of cloning, individual amplicons are incorporated into vectors. Each vector–insert (ideally the insert is an amplicon of interest) is then incorporated into a single bacterium, and the bacteria multiply each vector–insert pair. Many different strategies for cloning PCR amplicons have been described (Frohman, 1994; Sambrook and Russell, 2001). In most situations and laboratories, TA cloning (Holton and Graham, 1991; Marchuk *et al*., 1991; Sambrook and Russell, 2001) will be easiest to implement. Several vendors (e.g., Invitrogen, Carlsbad, CA, and Promega, Madison, WI) offer products for such cloning experiments. In the example below, we describe some specific techniques that may be useful when working with amplicons from multigene families.

Following cloning, the individual bacterial colonies must be screened to ensure inserts of the appropriate size have been obtained. Various methods have been described to achieve this goal (Sambrook and Russell, 2001). In many labs, this goal is most easily accomplished by using PCR and primers that bind to the multiple cloning site of the vector (e.g., M13 forward and M13 reverse). The number of colonies that should be screened depends on the efficiency of cloning products of the appropriate size and the number that are needed for sequencing (see below).

Example: Feather β*-Keratins from Individual Birds (Fig. 1).* To reduce cloning small nonspecific PCR amplicons (i.e., primer dimers), PCR products were precipitated by adding an equal volume of 20% PEG with 2.5 *M* NaCl, centrifuging, washed twice with 80% EtOH (Kusukawa, 1990), and quantified on a 1.5% agarose gel using varying quantities of

lambda DNA as a standard. To improve ligation efficiency, A-tailing reactions were carried out for 30 min at 70° in 10-μl volumes with final concentrations of 50 m*M* KCl, 10 m*M* Tris–HCl pH 9, 1% Triton X-100, 2.0 m*M* $MgCl_2$, 0.2 m*M* dATP, 2 units of *Taq* DNA Polymerase (Promega, Madison, WI), and 250 ng of precipitated PCR amplicons. Three microliters (75 ng) of the A-tailing reaction was used in the ligation reaction for each sample. Ligations were performed using pGEM-T Vector System kits (Promega, Madison, WI) according to the manufacturer's protocol. Following ligation, 2 volumes (20 μl) of TE were added to each reaction, and each sample was incubated at 65° for 20 min to dissociate the ligase from the DNA. Transformations were performed using Epicurian Coli XL-10 Gold™ Ultracompetent Cells (Stratagene, La Jolla, CA) according to the manufacturer's protocol or by electroporation of this strain.

For each sample, 48–192 colonies, identified as potentially positive for inserts by color screening, were picked and transferred to 40 μl of water. These colony suspensions were used as templates to determine insert size by performing PCR using M13 Forward and Reverse primers (Stratagene, La Jolla, CA) and an initial denaturation period of 10 min. PCR amplicons were examined on 1.5% agarose gels (Sambrook and Russell, 2001); those that contained inserts of 800–1100 bp were used as templates for sequencing reactions.

Automated DNA Sequencing of Clones

Following identification of clones with inserts of interest, the DNA sequence needs to be determined. Various options are available (Sambrook and Russell, 2001). Researchers should choose the method that will be most efficient and economical for them. Most major research institutions have automated DNA sequencers available in core labs and/or in the labs of genomics researchers. The newest generation of capillary sequencers allow read lengths of 1000 bases or more and run costs less than $1/lane. Outsourcing to commercial or core labs at other institutions that have these capillary instruments is a viable option, even when other automated sequencers are available at a researcher's own institution. Because it is straightforward to complete cycle sequencing reactions in any lab equipped to conduct PCR, many researchers will find it economical to do the reactions themselves and simply employ the commercial/core lab to run reactions on their instruments, sending the unedited chromatograms back to the researchers.

After the chromatograms are generated, they must be edited to ensure accuracy of the sequence obtained. A variety of commercial (Sequencher–Genecodes, Ann Arbor, MI; Lasergene–DNAStar, Inc., Madison, WI;

GeneTool–BioTools, Inc., Edmonton, Alberta, Canada; e-Seq–Licor, Lincoln, NE; Vector NTI–InforMax, Frederick, MD) and free (e.g., Sequence Viewer–Applied Biosystems, Foster City, CA; AutoEditor–The Institute for Genomic Research, Rockville, MD) software packages are available for this purpose. Many research institutions support free software and shared user licenses of the commercial software, often through the core sequencing facilities. Researchers should invest sufficient resources in determining the accuracy of DNA sequences to justify the considerable expenses incurred by subsequent steps of this approach.

Example: Sequencing Clones of Feather β*-Keratins (Fig. 1).* Two methods of automated DNA sequencing were used. To produce clean reads of more than 1000 bases, a LICOR 4000L (Lincoln, NE) automated sequencer with fluorescently labeled primers was used. Sequences from both DNA strands were determined from the products of the colony screening PCRs, which had been purified by PEG precipitation and quantified. Cycle sequencing (Amersham, Cleveland, OH) of purified products was conducted with M13 Forward and Reverse primers labeled with IRD41 (LICOR, Lincoln, NE). Cycle sequencing parameters were 95° for 2 min, followed by 25 cycles of 95° for 30 s, 55° for 30 s, and 70° for 1.25 min.

The second dataset was generated using an ABI 377XL automated sequencer (Applied Biosystems 1998), primer R972, and an internal primer F433 (Table I) designed specifically for the purpose of allowing us to focus on the protein coding region of the gene. Sequences from both DNA strands were determined directly from the products of the colony PCRs using Big-Dye terminator chemistry (Applied Biosystems, Foster City, CA). All cycle sequencing reactions were carried out in 10-μl volumes with 0.32 μM primer, about 50 ng of PCR product, and Big-Dye terminators using the manufacturer's specifications except that the terminator mix was diluted 1:1 with halfBD (GENPAK Ltd., Stony Brook, NY) or a homemade equivalent (400 mM Tris–HCl pH 9, 10 mM $MgCl_2$; see http://www.genome.ou.edu/proto.html for details).

Chromatograms of forward and reverse sequences from each sample were imported into Sequencher 3.1.1 (Gene Codes, Ann Arbor, MI), which was used to edit and align the individual sequences into contigs. Sequencher, using the universal translation codes, allowed inference of the amino acid sequence, which was then exported for further manipulation.

Aligning Amino Acid Sequences and Identifying Domains of Interest

After amino acid sequences of interest have been obtained (from steps 1 and 5), they need to be aligned. As with aligning DNA sequences, several software tools and protocols are available to assist with aligning amino acid

TABLE I

ALIGNMENT OF AMINO ACID SEQUENCES FOR β-KERATINS WAS DETERMINED DIRECTLY OR INFERRED FROM DNA SEQUENCES[a]

	Amino acid position									
	0	1	2	3	4	5	6	7	8	8
	1	0	0	0	0	0	0	0	0	3
Turkey	MSCYDLCRPC	----	GPTPLANSCNEPCVRQCQDSRV							VIEPSPVVVTLPGPILSSFPQNTAVGSTTSAAVGSILSEEGVPINSG
vulture										
Wood stork					I					A.....S...
Pigeon			NP.L...QC				Q			S.....C.....
Chicken							Q			S.....S...
FBK A										
Chicken			FHV.Q	---	....	-	T...Q			T...SA.....A..AG.....S...
feather-like										
Chicken			SSLCA.ACVAT	-	.....	D.....	P..T...Q.PAT...F			YA.....AGVP.....GMGGTFGRGAGF
claw										
Chicken			P.TSCISR.Q.I.D.G		.....	P..TT..Q.P				.....F.....DSV....SGAPIF.GSSLGY.GSSLGY
scale										

Note: The feather β-keratin antiserum (anti-FβK) was generated using a synthetic polypeptide whose sequence was identical to the underlined amino acids of the turkey vulture sequence.

[a] Sequences correspond (from top to bottom) to accessions: AF308826, AF308827, KRPYF4, X17511, X17521, M37698, and X00315 of GenBank or GenPept.

sequences (http://us.expasy.org/tools/). The amount of effort required for amino acid alignment will likely range from nearly trivial to extraordinarily difficult. The goal is for the alignment to be of sufficient quality that regions of interest can be identified. In the overall approach we are describing, a region of interest is a string of amino acid sequence that is conserved within a subgroup of multigene family members, but variable among subgroups. Thus, researchers should choose a method that produces an alignment of sufficient quality to achieve that goal.

Example: The β*-Keratin Multigene Family (Fig. 1).* Amino acid sequences from GenBank and inferred from DNA sequences (described earlier) were obtained. We used the Clustal algorithm implemented in Sequence Navigator 1.0.1 (Applied Biosystems, Foster City, CA) to align the sequences. From this alignment (Table I), it is clear that all β-keratins share regions of high similarity (e.g., residues 44–57 have only one variable amino acid among all β-keratins), but regions of dissimilarity exist among the subgroups (e.g., residues 61–83 have few substitutions within FβKs, but several to many substitutions compared with other subgroups). Using information from an alignment that included many more FβK sequences, we chose a region of interest to synthesize as an oligopeptide.

Synthesizing an Oligopeptide

One of the advantages of knowing the specific sequences of families of related genes is that regions of high similarity within a subfamily and heterogeneity among subfamilies can be ascertained. However, other regions of high similarity among all members of the multigene family often exist within the same proteins. Thus, regions of varying similarity to other proteins are linked in any intact protein. If researchers rely on using intact proteins isolated from biological tissues to produce polyclonal antibodies (as is traditionally done [see step 8 below]), then the antibodies will recognize all proteins with similarity to any of the domains. Such polyclonal antibodies are, therefore, not specific to subfamilies (e.g., FβKs vs. ScβKs).

A strategy to avoid the problems associated with using intact proteins isolated from biological tissue is to synthesize an oligopeptide that is homologous to unique regions of interest. This strategy allows antibodies (step 8) to be made, which will only react with proteins containing the unique region of interest, even though other regions may have amino acid sequences homologous to other proteins.

Example: The β*-Keratin Multigene Family (Fig. 1).* Alignment of amino acid sequences for β-keratins revealed distinct amino acid sequences of FβKs, which were not present in other subfamilies. We used this unique

amino acid sequence, conserved in all FβKs, to synthesize a feather-specific peptide useful in the production of monospecific antiserum (Sawyer *et al*., 2003a). Based on sequence analysis and alignment of feather β-keratin amino acid sequences, the 23-mer VGSTTSAAVGSILSEEGVPINSG-CONH2 (a C-terminal amide) was synthesized on an applied Biosystems Pioneer automated peptide synthesizer according to the methods described in Sawyer *et al*. (2003a).

Producing an Antiserum

There are two convenient approaches to making high-specificity antibodies. One is through the technically demanding methods of producing hybridomas following fusion of activated murine lymphocytes and immortalized myeloma cells (Zola, 1987), and the other is through the use of oligopeptides in conventional methods involving injection of these antigens into mammals (O'Guin *et al*., 1982; Sawyer *et al*., 2003a). Monoclonal antibodies produced by hybridomas are useful but often tend to be less efficacious when epitopes in the original proteins are rare or conformationally cryptic. Oligopeptides used for injection into a mammal, on the other hand, have multiple but limited epitopes that are more likely available in the intact protein under a wide range of conditions (e.g., immunoblotting of proteins and immunostaining of cells and tissues). Antisera made in animals are much less technically demanding to produce because of the relative simplicity of animal use methods (O'Guin *et al*., 1982).

Example: The β*-Keratin Multigene Family (Fig. 1).* In the case of our work, a highly specific (FβK antiserum was produced in a male New Zealand white rabbit (Sawyer *et al*., 2003a). The synthetic peptide antigen was cross-linked to keyhole limpet hemocyanin (KLH) using glutaraldehyde (Sawyer *et al*., 2003a). The primary and secondary injections of dialyzed Freund's emulsified antigen contained 200 mg of the synthetic peptide cross-linked with KLH. Serum was collected and processed 14 days after the second injection. Initial screening of the serum against extracts of feather keratins demonstrated specificity for the feather β-keratins. Preimmune serum controls were negative. For convenience, the antiserum was divided into 1-ml aliquots and frozen at -70° for preservation and future use.

Step 8a: Immunoblotting (Western Blots)

To verify that the synthetic peptide did indeed present a unique set of epitopes for the production of an antiserum that would react specifically against FβKs, we analyzed its reactivity against the epidermal polypeptides extracted from feather and scale tissues (O'Guin *et al*., 1982; Sawyer

et al., 2003a). The epidermal keratins were extracted with Triton X-100 in 1.5 *M* KCl and separated by electrophoresis on a 10% polyacrylamide gel electrophoresis (PAGE). The peptides were stained with Coomassie brilliant blue (Fig. 3). Note that the α-keratins, which are also prevalent in epidermal tissues of all chordates (O'Guin and Sawyer, 1982; Sawyer *et al*., 2000; Whitbread *et al*., 1991) migrate as bands in the molecular weight range of 40–70 kd, while scale β-keratins migrate in the range of 17–20 kd. The FβKs migrate slightly lower at 10–14 kd (Haake *et al*., 1984; O'Guin and Sawyer, 1982; Whitbread *et al*., 1991).

Keratins separated by PAGE were transferred to nitrocellulose or *Immobulon* membranes (Sawyer *et al*., 2003a) and incubated overnight in a 1:5000 dilution of anti-FβK in phosphate-buffered saline. Membranes were rinsed thoroughly and exposed to secondary antibodies (goat anti-rabbit IgG linked to horseradish peroxidase). An appropriate substrate

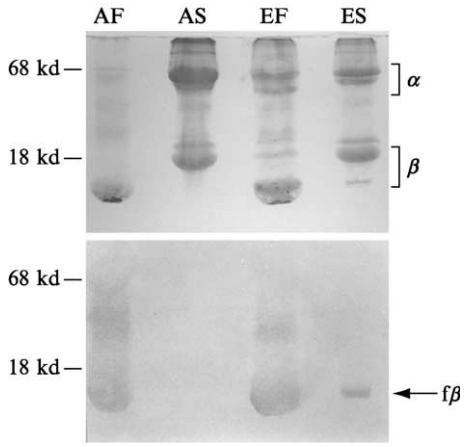


FIG. 3. The upper panel shows the epidermal keratins of adult feather (AF), adult scutate scale (AS), embryonic feather (EF), and embryonic scutate scale (ES) extracted with Triton X-100/1.5*M* KCl and separated by electrophoresis on a 10% polyacrylamide gel. The keratins are stained with Coomassie brilliant blue. The alpha (α) keratins migrate as bands in the range of 40–70 kd, whereas the scale-type β keratins (ScβKs) migrate to the range of 17–20 kd, and the feather-type β keratins (FβKs) migrate to the range of 10–14 kd. The lower panel shows the corresponding Western blot. The keratins separated as shown in the upper panel were transferred to a nitrocellulose membrane and incubated overnight with the FβK antiserum, and the FβK antiserum was localized with goat anti-rabbit horseradish peroxidase. The FβK is present in the AFs and EFs as evidenced by the broad band in the molecular weight range of the β keratins. The bands at higher molecular weight are the result of aggregation known to occur with β keratins. The FβK band is absent from the adult scale (AS) but is present in the embryonic scale (ES). The FβK antiserum does not react with the ScβKs present in both the adult and the ES (after Sawyer *et al.* [2003]).

(e.g., diaminobenzidine) was used to produce a precipitate on the membrane to mark the location of antigens (O'Guin *et al*., 1982; Sawyer *et al*., 2003a). Figure 3 shows the localization of FβK antiserum on extracts of epidermal β keratins of adult feather, adult scale, embryonic feather and embryonic scale.

Step 8b: Immunostaining of Histological Sections

As important as the specificity of this antiserum is in determining the biochemical identity of the FβKs in immunoblots, it is equally important in functional studies to determine the timing of expression and particularly the spatial distribution of the β-keratins in the cells of the structures themselves as they develop. For this aspect of the analysis, we employed immunohistochemistry of tissues that were embedded in paraffin and sectioned on a microtome (Sawyer *et al*., 2003a). The antiserum was used as a probe for location of the protein in actual feather and scale epidermal tissue. Specifically, we used indirect immunofluorescence techniques in conjunction with confocal microscopy (Sawyer *et al*., 2003a). The advantage of confocal microscopy is in the resolution attainable because of a dramatic reduction in flare and the ability to optically section thick sections (10–20 μm) of tissue. Combining paraffin sectioning with confocal analysis allows the collection of data on serial sets of sections. In Fig. 4, confocal microscope images of feather and scale show localization of feather type β-keratins using FITC-tagged anti-rabbit IgG antiserum to localize the binding of the FβK antiserum to its specific antigens.

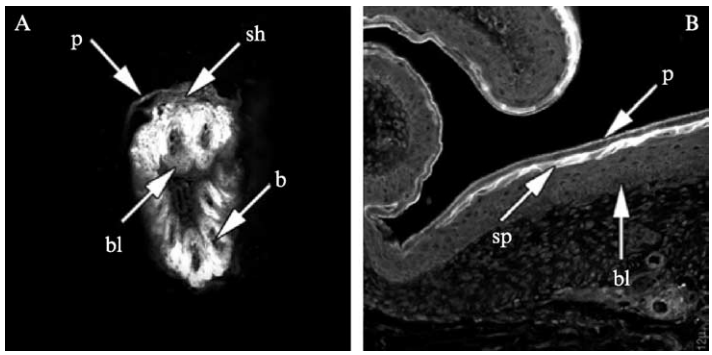


FIG. 4. (A) A confocal microscope image of a 17-day-old embryonic feather localizing feather-type β keratin using FITC-tagged anti-rabbit immunoglobulin G (IgG) antiserum, which localized the anti-FβK. (B) The location of the feather-type β keratins in a 17-day-old embryonic scutate scale. (After Sawyer *et al*. [2003].) b, barb ridge; sh, sheath cells; p, primary periderm cells; sp, subperidermal cells; bl, stratum basale region.

Unlike the results with the Western blots, which solely demonstrate reactivity to the electrophoretically separated proteins, the results of immunohistochemistry indicate where such proteins reside in the cells and tissues themselves. For example, in Fig. 4 individual cells of the barb ridges of an embryonic day 17 feather are easily defined by both position and intensity. In Fig. 4B, an embryonic day 17 scutate scale (found on the anterior surface of the bird's foot) displays two separate and distinct cell layers in which the FβK are localized. This approach adds a new dimension to the functional genomic analysis presented here, in that for the first time it can be unambiguously demonstrated that embryonic scutate scales express FβKs as demonstrated by both Western blots and immunohistochemistry.

Concluding Remarks

There are clear advantages to understanding developmental processes in light of a relatively complete sequence of events from transcription to translation to differentiation and histological organization. One advantage is that we are not left with much doubt about what the fate of such expression is in terms of functionality. Unlike many systems in which the expression of an enzyme or transcription factor is linked to many different pathways or is transient in presentation, the use of structural elements such as β-keratin has a fate that is less complicated. Another advantage for functional genomics is the ability to compare species with historical evolutionary relationships and determine what is unique about processes such as morphogenesis and those aspects of the process that are shared. With respect to feather development, it is now becoming clear that there are events in the morphogenesis of feathers that relate not only to birds, but also potentially to modern reptiles and their predecessors. The functional genomics approach employed in this methods chapter embraces the desire to uncover the common themes of evolutionary developmental biology while retaining our interest in the unique aspects of each species under study.

References

Alibardi, L., and Sawyer, R. H. (2002). Immunocytochemical analysis of beta (β) keratin in the epidermis of chelonians, lepidosaurians and archosaurians. *J. Exp. Zool.* **293,** 27–38.

Carver, W. E., and Sawyer, R. H. (1987). Development and keratinization of the epidermis in the common lizard, *Anolis carolinensis*. *J. Exp. Zool.* **243,** 435–443.

French, J. O. (2001). ''Characterization and evolutionary relationships among copies of feather β-keratin genes.'' M.S. Thesis, University of South Carolina, Columbia.

Frohman, M. A. (1994). Cloning PCR products. *In* ''The Polymerase Chain Reaction'' (K. B. Mullis, F. Ferre, and R. A. Gibbs, eds.), pp. 14–37. Birkhauser, Boston, MA.

Haake, A. R., Koenig, G., and Sawyer, R. H. (1984). Avian feather development: Relationships between morphogenesis and keratinization. *Dev. Biol.* **106,** 406–413.

Hall, B. G. (2001). ''Phylogenetic Trees Made Easy: A How-To Manual for Molecular Biologists. Sinauer Associates, Inc., Sunderland, MA.

Holton, T. A., and Graham, M. W. (1991). A simple and efficient method for direct cloning of PCR product using ddT-tailed vectors. *Nucleic Acids Res.* **19,** 1156.

Inglis, A. S., Gillespie, J. M., Roxburgh, C. M., Whittaker, L. A., and Casagranda, F. (1987). *In* ''Proteins, Structure, and Function'' (J. L. L'Italien, ed.), p. 757. Plenum Press, New York and London.

Kemp, D. J., and Rogers, G. E. (1972). Differentiation of avian keratinocytes: Characterization and relationship of the keratin proteins of adult and embryonic feather and scales. *Biochem.* **11,** 969–975.

Kusukawa, N., Uemori, T., Asada, K., and Kato, I. (1990). Rapid, reliable protocol for direct sequencing of material amplified by the PCR. *Biotechniques* **9,** 66–72.

Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1991). Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19,** 1154.

Molloy, P. L., Powell, B. C., Gregg, K., Barone, E. D., and Rogers, G. E. (1982). Organization of feather keratin genes in the chick genome. *Nucleic Acids Res.* **10,** 6007–6021.

O'Guin, W. M., Knapp, L. W., and Sawyer, R. H. (1982). Biochemical and immunohistochemical localization of alpha and beta keratin in avian scutate scale. *J. Exp. Zool.* **220,** 371–376.

O'Guin, W. M., and Sawyer, R. H. (1982). Avian scale development: VIII. relationship between morphogenetic and biosynthetic differentiation. *Dev. Biol.* **89,** 485–492.

Palumbi, S. R. (1996). Nucleic acids III: The polymerase chain reaction. *In* ''Molecular Systematics'' (D. M. Hillis, C. Moritz, and B. K. Mable, eds.), 2 Ed., pp. 205–247. Sinauer Associates, Inc., Sunderland, MA.

Presland, R. B., Gregg, K., Molloy, P. L., Morris, C. P., Crocker, L. A., and Rogers, G. E. (1989a). Avian keratin genes. I. A molecular analysis of the structure and expression of a group of feather keratin genes. *J. Mol. Biol.* **209,** 549–559.

Presland, R. B., Whitbread, L. A., and Rogers, G. E. (1989b). Avian keratin genes. II. chromosomal arrangement and close linkage of three gene families. *J. Mol. Biol.* **209,** 561–576.

Rogers, G. E., Dunn, S., and Powell, B. (1998). Late events and the regulation of keratinocyte differentiation in hair and feather follicles. *In* ''Molecular Basis of Epithelial Appendage Morphogenesis,'' pp. 315–340. R. G. Landes Company, Austin, TX.

Sambrook, J., and Russell, D. W. (2001). ''Molecular Cloning: A Laboratory Manual.'' Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sawyer, R. H., Glenn, T. C., French, J. O., Mays, B., Shames, R. B., Barnes, G. L., Rhodes, W., and Ishikawa, Y. (2000). The expression of beta (β) keratins in the epidermal appendages of reptiles and birds. *Am. Zool.* **40,** 530–539. Wiley-Liss, New Jersey.

Sawyer, R. H., and Knapp, L. W. (2003). Avian skin development and the origin of feathers. *In* ''Molecular and Developmental Evolution'' (C.-M. Chuong, ed.), pp. 57–72.

Sawyer, R. H., Salvatore, B. A., Potylicki, T.-T. F., French, J. O., Glenn, T. C., and Knapp, L. W. (2003a). Origin of feathers: Feather beta (β) keratins are expressed in discrete epidermal cell populations of embryonic scutate scales. *J. Exp. Zool. (Mol. Dev. Evol.)* **295B,** 12–24.

Sawyer, R. H., Washington, L. D., Salvatore, B. A., Glenn, T. C., and Knapp, L. W. (2003b). Origin of archosaurian integumentary appendages: The bristles of the wild turkey beard express feather-type β keratins. *J. Exp. Zool. (Mol. Dev. Evol.)* **297B,** 1–8.

Shames, R. B., Knapp, L. W., Carver, W. E., and Sawyer, R. H. (1988). Identification, expression and localization of beta keratin gene products during development of avian scutate scales. *Differentiation* **38,** 115–123.

Thompson, J. D., Gibson, T. J., and Higgins, D. G. (1994). The ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position, specific gap penalties, and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The ClustalX-Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25,** 4876–4882.
Walker, I. D., and Bridgen, J. (1976). The keratin chains of avian scale tissue, sequence heterogeneity and the number of scale keratin genes. *Eur. J. Biochem.* **67,** 283–293.
Whitbread, L. A., Gregg, K., and Rogers, G. E. (1991). The structure and expression of a gene encoding chick claw keratin. *Gene* **101,** 223–229.
Wilton, S. D., Crocker, L. A., and Rogers, G. E. (1985). Isolation and characterization of keratin m RNA from the scale epidermis of the embryonic chick. *Biochim. Biophys. Acta* **824,** 201–201.
Zola, H. (1987). ''Monoclonal Antibodies: A Manual of Techniques.'' CRC Press, Boca Raton, FL.

[34] Applications of Ancestral Protein Reconstruction in Understanding Protein Function: GFP-Like Proteins

By BELINDA S. W. CHANG, JUAN A. UGALDE, and MIKHAIL V. MATZ

Abstract

Recreating ancestral proteins in the laboratory increasingly is being used to study the evolutionary history of protein function. More efficient gene synthesis techniques and the decreasing costs of commercial oligo-synthesis are making this approach both simpler and less expensive to perform. Developments in ancestral reconstruction methods, particularly more realistic likelihood models of molecular evolution, allow for the accurate reconstruction of more ancient proteins than previously possible. This chapter reviews phylogenetic methods of ancestral inference, strategies for investigating alternative reconstructions, gene synthesis, and design, and an application of these methods to the reconstruction of an ancestor in the green fluorescent protein family.

Introduction

Ancestral protein reconstruction allows for the recreation of protein evolution in the laboratory so that it can be studied directly. This approach is a natural extension of experimental studies that examine present-day

0076-6879/05 $35.00